# Full wwPDB X-ray Structure Validation Report ⓘ

Oct 5, 2023 – 05:18 AM EDT

| | | |
|---|---|---|
| PDB ID | : | 6VJS |
| Title | : | Escherichia coli RNA polymerase and ureidothiophene-2-carboxylic acid complex |
| Authors | : | Murakami, K.S.; Molodtsov, V. |
| Deposited on | : | 2020-01-17 |
| Resolution | : | 4.02 Å(reported) |

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references ⓘ) were used in the production of this report:

| | | |
|---|---|---|
| MolProbity | : | FAILED |
| Mogul | : | 1.8.5 (274361), CSD as541be (2020) |
| Xtriage (Phenix) | : | 1.13 |
| EDS | : | FAILED |
| buster-report | : | 1.1.7 (2018) |
| Percentile statistics | : | 20191225.v01 (using entries in the PDB archive December 25th 2019) |
| Ideal geometry (proteins) | : | Engh & Huber (2001) |
| Ideal geometry (DNA, RNA) | : | Parkinson et al. (1996) |
| Validation Pipeline (wwPDB-VP) | : | 2.35.1 |

# 1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
*X-RAY DIFFRACTION*

The reported resolution of this entry is 4.02 Å.

There are no overall percentile quality scores available for this entry.

MolProbity and EDS failed to run properly - the sequence quality summary graphics cannot be shown.

# 2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 56011 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called DNA-directed RNA polymerase subunit alpha.

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | A | 234 | Total | C | N | O | S | 0 | 0 | 0 |
| | | | 1809 | 1126 | 319 | 357 | 7 | | | |
| 1 | B | 301 | Total | C | N | O | S | 0 | 0 | 0 |
| | | | 2337 | 1462 | 411 | 456 | 8 | | | |
| 1 | F | 234 | Total | C | N | O | S | 0 | 0 | 0 |
| | | | 1809 | 1126 | 319 | 357 | 7 | | | |
| 1 | G | 229 | Total | C | N | O | S | 0 | 0 | 0 |
| | | | 1775 | 1106 | 313 | 350 | 6 | | | |

- Molecule 2 is a protein called DNA-directed RNA polymerase subunit beta.

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|---|
| 2 | C | 1336 | Total | C | N | O | S | 0 | 0 | 0 |
| | | | 10540 | 6614 | 1836 | 2047 | 43 | | | |
| 2 | H | 1336 | Total | C | N | O | S | 0 | 0 | 0 |
| | | | 10540 | 6614 | 1836 | 2047 | 43 | | | |

- Molecule 3 is a protein called DNA-directed RNA polymerase subunit beta'.

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|---|
| 3 | D | 1164 | Total | C | N | O | S | 0 | 0 | 0 |
| | | | 9093 | 5717 | 1626 | 1704 | 46 | | | |
| 3 | I | 1160 | Total | C | N | O | S | 0 | 0 | 0 |
| | | | 9061 | 5698 | 1619 | 1698 | 46 | | | |

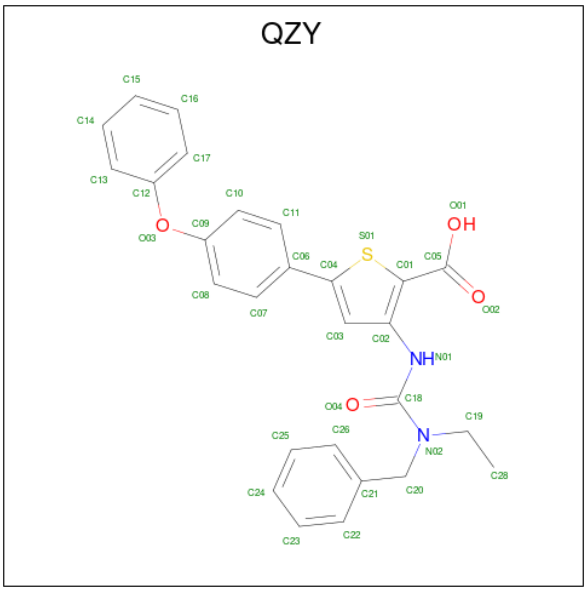
- Molecule 4 is a protein called DNA-directed RNA polymerase subunit omega.

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|---|
| 4 | E | 90 | Total | C | N | O | S | 0 | 0 | 0 |
| | | | 708 | 430 | 136 | 141 | 1 | | | |
| 4 | J | 76 | Total | C | N | O | S | 0 | 0 | 0 |
| | | | 605 | 368 | 115 | 121 | 1 | | | |

- Molecule 5 is a protein called RNA polymerase sigma factor RpoD.

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|---|
| 5 | X | 471 | Total | C | N | O | S | 0 | 0 | 0 |
| | | | 3832 | 2398 | 687 | 724 | 23 | | | |
| 5 | Y | 475 | Total | C | N | O | S | 0 | 0 | 0 |
| | | | 3862 | 2413 | 692 | 734 | 23 | | | |

- Molecule 6 is 3-{[benzyl(ethyl)carbamoyl]amino}-5-(4-phenoxyphenyl)thiophene-2-carboxylic acid (three-letter code: QZY) (formula: $C_{27}H_{24}N_2O_4S$) (labeled as "Ligand of Interest" by depositor).




| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf |
|---|---|---|---|---|---|---|---|---|---|
| 6 | D | 1 | Total | C | N | O | S | 0 | 0 |
| | | | 34 | 27 | 2 | 4 | 1 | | |

- Molecule 7 is ZINC ION (three-letter code: ZN) (formula: Zn).

| Mol | Chain | Residues | Atoms | | ZeroOcc | AltConf |
|---|---|---|---|---|---|---|
| 7 | D | 2 | Total | Zn | 0 | 0 |
| | | | 2 | 2 | | |
| 7 | I | 2 | Total | Zn | 0 | 0 |
| | | | 2 | 2 | | |

- Molecule 8 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

| Mol | Chain | Residues | Atoms | ZeroOcc | AltConf |
|---|---|---|---|---|---|
| 8 | D | 1 | Total Mg<br>1 1 | 0 | 0 |
| 8 | I | 1 | Total Mg<br>1 1 | 0 | 0 |

MolProbity and EDS failed to run properly - this section is therefore empty.

# 3 Data and refinement statistics (i)

EDS failed to run properly - this section is therefore incomplete.

| Property | Value | Source |
|---|---|---|
| Space group | P 21 21 21 | Depositor |
| Cell constants<br>a, b, c, $\alpha$, $\beta$, $\gamma$ | 187.08Å 205.50Å 309.88Å<br>90.00° 90.00° 90.00° | Depositor |
| Resolution (Å) | 50.09 – 4.02 | Depositor |
| % Data completeness<br>(in resolution range) | 99.3 (50.09-4.02) | Depositor |
| $R_{merge}$ | 0.11 | Depositor |
| $R_{sym}$ | (Not available) | Depositor |
| $< I/\sigma(I) >$ [1] | 1.97 (at 4.00Å) | Xtriage |
| Refinement program | PHENIX 1.17.1_3660 | Depositor |
| R, $R_{free}$ | 0.230 , 0.276 | Depositor |
| Wilson B-factor (Å$^2$) | 157.4 | Xtriage |
| Anisotropy | 0.370 | Xtriage |
| L-test for twinning[2] | $< \|L\| > = 0.47$, $< L^2 > = 0.30$ | Xtriage |
| Estimated twinning fraction | No twinning to report. | Xtriage |
| Total number of atoms | 56011 | wwPDB-VP |
| Average B, all atoms (Å$^2$) | 205.0 | wwPDB-VP |

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.64% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

# 4 Model quality

## 4.1 Standard geometry

MolProbity failed to run properly - this section is therefore empty.

## 4.2 Too-close contacts

MolProbity failed to run properly - this section is therefore empty.

## 4.3 Torsion angles

### 4.3.1 Protein backbone

MolProbity failed to run properly - this section is therefore empty.

### 4.3.2 Protein sidechains

MolProbity failed to run properly - this section is therefore empty.

### 4.3.3 RNA

MolProbity failed to run properly - this section is therefore empty.

## 4.4 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

## 4.5 Carbohydrates

There are no monosaccharides in this entry.

## 4.6 Ligand geometry

Of 7 ligands modelled in this entry, 6 are monoatomic - leaving 1 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond

length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

| Mol | Type | Chain | Res | Link | Bond lengths | | | Bond angles | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Counts | RMSZ | $\#\|Z\| > 2$ | Counts | RMSZ | $\#\|Z\| > 2$ |
| 6 | QZY | D | 1501 | - | 33,37,37 | 4.09 | 5 (15%) | 38,50,50 | 1.79 | 7 (18%) |

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

| Mol | Type | Chain | Res | Link | Chirals | Torsions | Rings |
|---|---|---|---|---|---|---|---|
| 6 | QZY | D | 1501 | - | - | 4/20/26/26 | 0/4/4/4 |

All (5) bond length outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed(Å) | Ideal(Å) |
|---|---|---|---|---|---|---|---|
| 6 | D | 1501 | QZY | C03-C02 | 21.50 | 1.59 | 1.39 |
| 6 | D | 1501 | QZY | C03-C04 | 5.19 | 1.44 | 1.37 |
| 6 | D | 1501 | QZY | C18-N02 | 4.83 | 1.45 | 1.36 |
| 6 | D | 1501 | QZY | C18-N01 | 4.56 | 1.45 | 1.37 |
| 6 | D | 1501 | QZY | O04-C18 | -2.22 | 1.19 | 1.23 |

All (7) bond angle outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed($^o$) | Ideal($^o$) |
|---|---|---|---|---|---|---|---|
| 6 | D | 1501 | QZY | C03-C02-N01 | -6.10 | 122.34 | 130.12 |
| 6 | D | 1501 | QZY | N01-C18-N02 | 4.57 | 121.12 | 115.89 |
| 6 | D | 1501 | QZY | C02-N01-C18 | -4.37 | 114.21 | 125.52 |
| 6 | D | 1501 | QZY | C09-O03-C12 | -3.24 | 111.23 | 118.80 |
| 6 | D | 1501 | QZY | C20-N02-C19 | 3.15 | 120.22 | 116.54 |
| 6 | D | 1501 | QZY | O01-C05-O02 | 2.84 | 129.65 | 123.35 |
| 6 | D | 1501 | QZY | O04-C18-N01 | -2.19 | 118.77 | 123.61 |

There are no chirality outliers.

All (4) torsion outliers are listed below:

| Mol | Chain | Res | Type | Atoms |
|---|---|---|---|---|
| 6 | D | 1501 | QZY | C03-C04-C06-C11 |
| 6 | D | 1501 | QZY | C03-C04-C06-C07 |
| 6 | D | 1501 | QZY | C28-C19-N02-C18 |

*Continued on next page...*

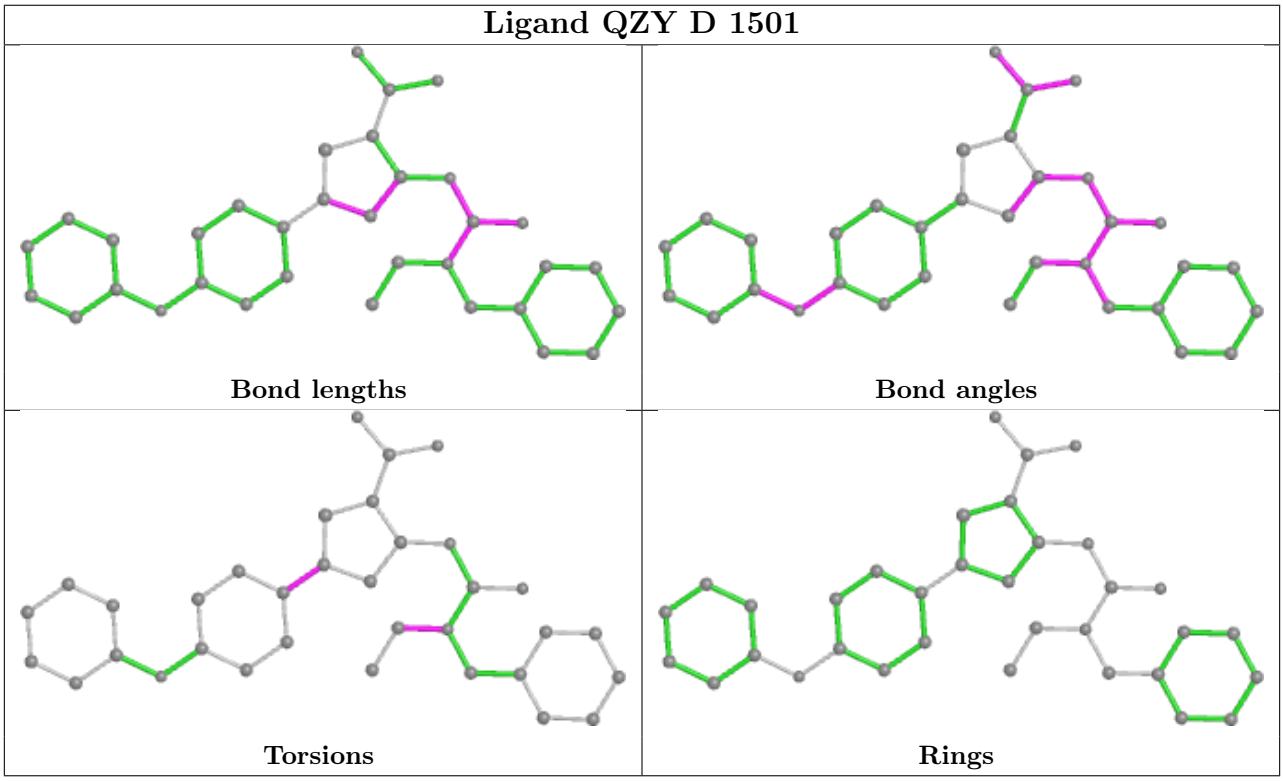
*Continued from previous page...*

| Mol | Chain | Res | Type | Atoms |
|---|---|---|---|---|
| 6 | D | 1501 | QZY | C28-C19-N02-C20 |

There are no ring outliers.

No monomer is involved in short contacts.

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.




## 4.7 Other polymers (i)

There are no such residues in this entry.

## 4.8 Polymer linkage issues (i)

The following chains have linkage breaks:

| Mol | Chain | Number of breaks |
|---|---|---|
| 1 | F | 1 |

All chain breaks are listed below:

| Model | Chain | Residue-1 | Atom-1 | Residue-2 | Atom-2 | Distance (Å) |
|---|---|---|---|---|---|---|
| 1 | F | 2:GLN | C | 3:GLY | N | 3.61 |

# 5 Fit of model and data

## 5.1 Protein, DNA and RNA chains

EDS failed to run properly - this section is therefore empty.

## 5.2 Non-standard residues in protein, DNA, RNA chains

EDS failed to run properly - this section is therefore empty.

## 5.3 Carbohydrates

EDS failed to run properly - this section is therefore empty.

## 5.4 Ligands

EDS failed to run properly - this section is therefore empty.

## 5.5 Other polymers

EDS failed to run properly - this section is therefore empty.